TECHNICAL MEMORANDUM

STATISTICAL ANALYSIS OF LEAD SAMPLES COLLECTED FROM PROPERTIES LOCATED WITHIN A QUARTER MILE OF THE HERCULANEUM LEAD SMELTER BOUNDARY

Tetra Tech EM Inc. (Tetra Tech) was tasked by the U.S. Environmental Protection Agency (EPA) Region 7 Enforcement Fund Lead Removal program to conduct a trend analysis of soil lead concentrations at selected locations within Herculaneum, Missouri (City). Specifically, the Tetra Tech Superfund Technical Assessment and Response Team (START) 2 was requested to review and analyze data that will enable EPA to determine if soil lead concentrations are increasing and the potential rate of increase at locations within 0.25 mile of the lead smelter boundary. The assessment was conducted under the authority of the Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA) and the Superfund Amendments and Reauthorization Act of 1986. The project was assigned under START Contract No. 68-S7-01-41, Task Order No. 0027.

This analysis was conducted using data collected from August 23, 2002 (Round 7) to December 22, 2003 (Round 15) from properties located within 0.25 mile of the smelter (house numbers 5, 6, 20, 22, and 24). This analysis had the following objectives:

(1) Evaluate differences in lead concentrations between front and backyards, and determine the most appropriate way to use data collected from individual quadrants in any statistical analysis of temporal trends.

(2) Conduct a detailed trend analysis for both individual properties and all properties combined, and provide estimates of the potential rates of change in lead concentrations, as well as an assessment of uncertainties associated with these estimates.

Objective 1:

An assessment of lead concentrations in front versus backyards was conducted using qualitative (graphical) and quantitative (paired-difference statistical tests) methods. Figure 1 presents outlier box plots (also referred to as box-and-whisker plots) of lead concentrations by sampling round for individual properties and for all properties combined in the left panel of each plot. In these plots, lead measurements from front yards are shown as open circles, and concentrations in backyards are shown as solid circles. The right-hand panel of each plot shows front versus backyard comparisons for all rounds combined. Horizontal lines that transect each box are the arithmetic mean concentrations. The bottom and top of each box represent the 25^{th} and 75^{th} percentiles of the data, respectively. The area between the 25^{th} and 75^{th} percentiles is referred to as the interquartile range (IQR). The horizontal line drawn across the

interior of the boxes represents the median or 50th percentile. The upper and lower bounds of the whiskers represent the highest and lowest values, respectively, that are not considered statistical outliers. Points falling above the whiskers are considered "high outliers," defined as values greater than the 75th percentile plus 1.5 times the IQR. "Low outliers" are defined as values less than the 25th percentile minus 1.5 times the IQR.

Statistical analysis of front and backyard concentrations was conducted using a paired-difference approach (Figure 2). This approach effectively evaluates the average difference in front versus backyard concentrations over all rounds of sampling. To conduct paired-difference testing, the average front and backyard concentrations (average of the two quadrants) are first calculated for each round of sampling. Next, the difference (front-backyard) between the two mean concentrations is calculated for each round. To conclude that no difference exists between front and backyard concentrations, the average difference over all sampling rounds should not differ statistically from zero, though some positive differences and some negative differences are expected. Figure 2 provides box plots, frequency histograms, and quantile tables of observed differences calculated for each sampling round. Figure 2 also provides results of both parametric (paired-difference t-test) and nonparametric (signed-rank test) tests of the two-sided null hypothesis (H_0) that the average difference is zero. If the probabilities associated with either of the two-sided tests (that is, "Prob > |t|" under the "Test Statistic" column in the embedded table) are less than or equal to 0.05 (5 percent), H_0 is rejected with the conclusion that lead concentrations are statistically different between front and backyards.

Conclusions for Objective 1. The statistical tests show that front yard concentrations exceed backyard concentrations for property numbers 22 and 24, and that backyard concentrations exceed front yard concentrations for property number 5. No statistically significant difference in concentrations was shown for property numbers 6 and 20. Examination of Figure 1 and results of the paired-difference tests suggest no consistent spatial pattern among properties when data are divided into front and backyard groupings. This is not surprising, as these groupings are not based on absolute orientations of front or backyards (or in effect, the relative degree of exposure of each area to air emissions of lead). A more technically defensible way to analyze the data at the scale of an individual quadrant (or pair of quadrants) would be to create groupings that reflect the relative degree of exposure to lead emissions. Under this approach, a front or backyard for an individual property would be grouped, as "exposed" or "relatively unexposed," and statistical analysis would proceed accordingly.

Objective 2:

Previous trend analyses conducted using a nonparametric test for monotonic trends (Mann-Kendall test) revealed significant increasing trends in lead concentrations for house numbers 5, 20, and 22 for samples collected from Round 7 through 14. To estimate potential rates of increase in lead concentrations, this analysis was repeated using linear regression analysis and additional data collected during Round 15. For this analysis, the earliest sampling date for Round 7 (August 23, 2002) was set at day = 0, and all subsequent sampling dates were converted to the number of days from this initial sampling date.

Figure 3 presents the results of regression analysis conducted for individual properties using two approaches. The first approach treats the data from individual quadrants as independent measurements. These results are presented as the set of plots at the left in Figure 3. Open and solid circles are used to represent data collected from front and backyards, respectively. The second approach calculates a median concentration for all four quadrants and regresses this against the time variable. As shown by the previous tests for monotonic trends, significant increasing trends appear for house numbers 5, 20, and 22. The trend results are significant for both the case in which individual measurements were used for each quadrant and the case in which only the median concentrations were used in the regression.

Figure 4 presents the results of regression analysis with all properties within 0.25 mile of the smelter combined. Three approaches are used in the analysis for the combined data:

(1) All measurements for all quadrants and houses are used in the regression of time (days) versus concentration.

(2) The medians for each round are calculated for individual houses, and the medians for all five houses are used in the regression of time (days) versus concentration.

(3) The grand median of all houses combined is calculated for each round, and the grand median for each round is used in the regression of time (days) versus concentration.

A significant increasing trend results from application of all three approaches using the data for all houses combined.

Uncertainty analysis. Graphical approaches as well as parametric and nonparametric statistical trend tests provide clear evidence of an increasing trend in lead concentrations for selected properties, and for the scenario in which all properties are treated as a single area. Therefore, estimating the potential rate at which lead concentrations may be increasing is of interest.

Estimating potential rates of increase in lead concentrations is not straightforward, and a number of uncertainties are associated with the types of trend analyses routinely conducted using environmental data sets. These uncertainties include:

1) Spatial variability in measured lead concentrations. That is, rates of increase are unlikely to be uniform over large areas, and different ways of grouping properties in this type of analysis are likely to yield different estimates. Moreover, different ways are available to treat replicate data for individual quadrants for each property. Figures 3 and 4 show clear differences in the relative spread of the data (that is, variability or noise) around the fitted regression line in cases where data for individual quadrants are included as independent measurements as opposed to representation by a measure of central tendency (median).

2) Selection of an appropriate statistical model for quantifying rates of increase. Linear models are the simplest to evaluate, but more complex non-linear models may better represent the true relationship between concentration and time.

3) Imprecision in measured concentrations associated with use of XRF analysis.

To better quantify the uncertainty associated with estimates of the rate of increase provided in Figures 3 and 4, two-sided 95-percent confidence limits were calculated for individual slope factors (regression coefficients) under several scenarios (Table 1). Two scenarios were considered. The first was a "worst case" scenario in which confidence limits were calculated for the property showing the most rapid increase in lead concentrations (house 20). The second was the scenario in which all properties were combined and treated as a single area. Regression coefficients (slope factors) in Table 1 are reported in units of milligrams per kilogram (mg/kg) per day as well as mg/kg per month. Confidence limits in Table 1 are expressed as monthly rates of increase.

It should be noted that EPA QA/G-9 (EPA 2000) provides several cautionary statements about applying regression analysis to environmental data. EPA (2000) notes that regression analysis is sensitive to extreme values (outliers), and that it is not well suited for handling censored data (that is, data below the detection limit). The validity of regression analysis also depends on two key assumptions: normally distributed errors and constant variance. EPA (2000) states that verifying these assumptions may be difficult or burdensome in routine practice, and suggests that regression analysis may be most useful as an "informal, quick, and easy screening tool for identifying strong linear trends." In the present analysis, the distribution of errors (that is, the deviation of individual measurements relative to the fitted regression line) was evaluated using residual plots. A plot of the residuals, or deviation of individual measurements from the "line of best fit," allows a rapid qualitative check of the assumption of normally distributed

errors. Residual plots appear at the bottom of the individual regression plots provided in Figures 3 and 4. Straightforward visual inspection of the residual plots does not provide evidence of serious departure from the assumption of normally distributed errors.

EPA (2000) also recommends a nonparametric approach for estimating slopes (Sen's Slope Estimate) as an alternative to linear regression. Calculation of Sen's Slope Estimate by hand is very tedious and not provided by most commercial statistical software packages. Sen's Slope Estimate was calculated in this analysis using a customized computer model similar to that used for running the Mann-Kendall test for monotonic trends. Calculations were performed for the worst-case scenario and the scenario in which the data for all houses are pooled. For both scenarios, only the data for the medians of all quadrants were used. These calculations yielded point estimates for the slope of 22 and 8 mg/kg per month for the worst-case and pooled scenarios, respectively.

Conclusions for Objective 2. The rates of increase shown in Table 1 should be viewed as rough estimates of the true rates of increase, given the uncertainties previously discussed. While regression analysis typically should not be used to predict future observations (that is, predictions should be limited to the range of the observed data), a coarse approximation of lead deposition over time could be made by assuming that the rate of increase predicted using the existing data would remain relatively constant. If so, the confidence limits calculated for the worst case scenario (increases as high as 6 to 11 mg/kg per month) and the scenario with all properties combined (increases as high as 1 to 6 mg/kg per month) could be used to estimate the time until average concentrations of lead for individual properties (or area-wide averages) will exceed the 400 mg/kg preliminary remediation goal (PRG) for lead. Collecting additional data over time and/or increasing sampling densities for individual rounds would increase the precision and accuracy of these estimates.

REFERENCE

U.S. Environmental Protection Agency (EPA). 2000. "Guidance for Data Quality Assessment: Practical Methods for Data Analysis: EPA QA/G-9, QA97 Version." EPA/600/R-96/084. Office of Research and Development. Washington, DC. July.

TABLE 1
RESULTS OF LINEAR REGRESSION ANALYSIS FOR SELECTED SCENARIOS

Scenario	Sample Size	Sampling Interval		Regression Coefficients for Days Versus Concentration			95 Percent Confidence Limits for Monthly Increase in Lead Concentrations	
		Initial Date	Final Date	Intercept	Slope[1]	S.E. (Slope)	Lower C.L.	Upper C.L.
Worst Case- House No. 20 (all quadrants treated as independent measurements)	32	8/26/2002	12/22/2003	72.74	0.28 (8.51)	0.04	6.12	10.80
Worst Case- House No. 20 (median of all quadrants used for each sampling round)	8	8/26/2002	12/22/2003	63.26	0.32 (9.53)	0.03	7.26	11.36
All Houses Combined (all quadrants treated as independent measurements)	148	8/23/2002	12/22/2003	85.98	0.11 (3.26)	0.03	1.70	4.81
All Houses Combined (medians for each house used for each round)	37	8/23/2002	12/22/2003	77.39	0.11 (3.43)	0.04	1.09	5.69
All Houses Combined (medians of all houses and quadrants used for each round)	9	8/23/2002	12/22/2003	68.74	0.12 (3.65)	0.02	2.03	5.00

Notes:

	Multiply the slope by 30 to get a point estimate of the monthly increase
C.L.	Confidence Limits
S.E.	Standard error of estimate
1	Numbers in parentheses are slope factors expressed as mg/kg per month

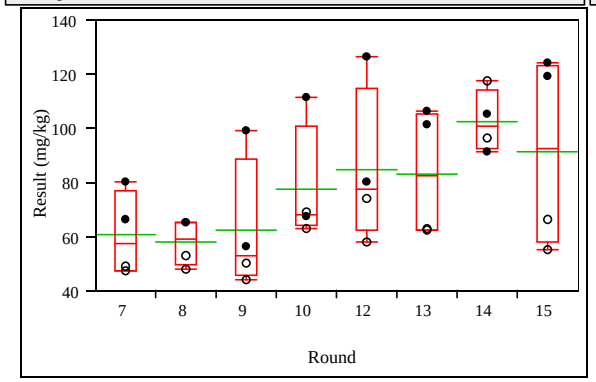
FIGURE 1

BOX PLOTS OF LEAD CONCENTRATIONS BY ROUND AND COMPARISON OF FRONT AND BACK YARDS FOR ALL ROUNDS COMBINED

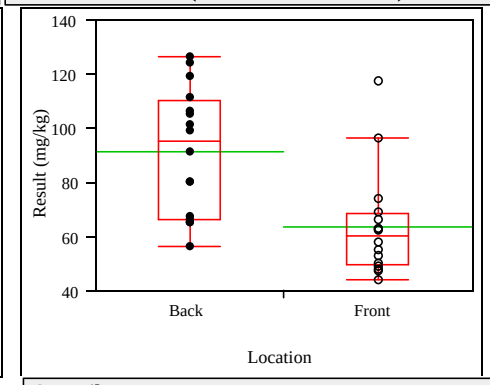
House=5

Open Circles (Front Yards) -- Solid Circles (Back Yards)

All Quadrants Plotted



Back Versus Front (All Rounds Combined)



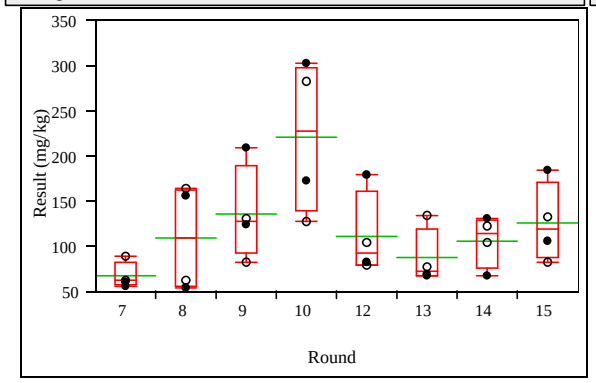
Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
Back	56	62.3	66.25	95	109.75	124.6	126
Front	44	46.1	49.25	60	68.25	102.3	117

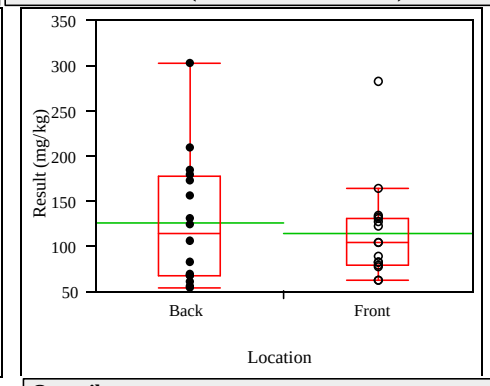
House=6

Open Circles (Front Yards) -- Solid Circles (Back Yards)

All Quadrants Plotted



Back Versus Front (All Rounds Combined)



Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
Back	53	54.4	66.25	114	177.25	236.2	302
Front	61	61.7	78.75	103.5	130.75	199.1	281

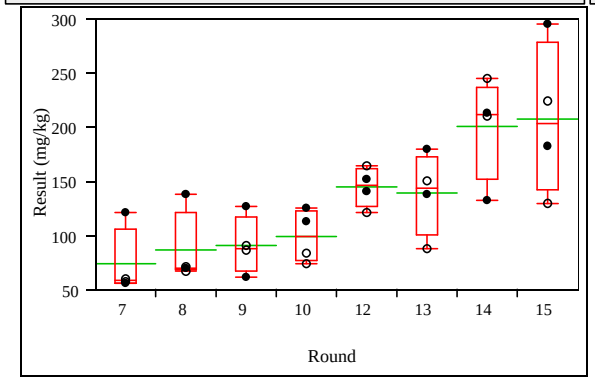
FIGURE 1 (CONTINUED)

BOX PLOTS OF LEAD CONCENTRATIONS BY ROUND AND COMPARISON OF FRONT AND BACK YARDS FOR ALL ROUNDS COMBINED

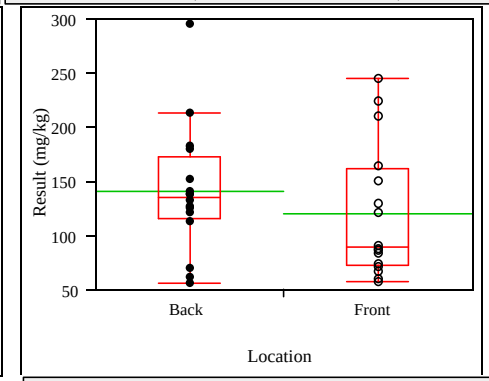
House=20

Open Circles (Front Yards) -- Solid Circles (Back Yards)

All Quadrants Plotted



Back Versus Front (All Rounds Combined)



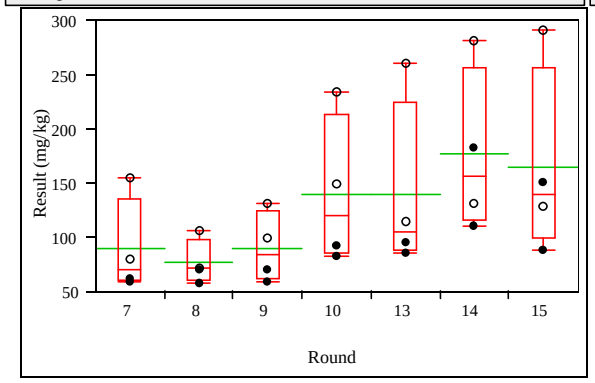
Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
Back	56	59.5	115	134.5	172.25	236.9	295
Front	57	59.1	71.75	89	160.5	230.3	245

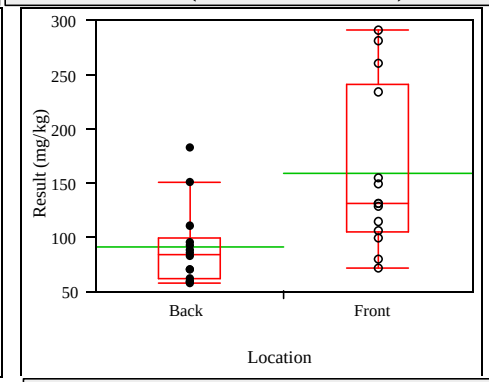
House=22

Open Circles (Front Yards) -- Solid Circles (Back Yards)

All Quadrants Plotted



Back Versus Front (All Rounds Combined)



Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
Back	57	58	60.5	83.5	98.75	166	182
Front	71	75	104.25	130.5	240.5	285.5	290

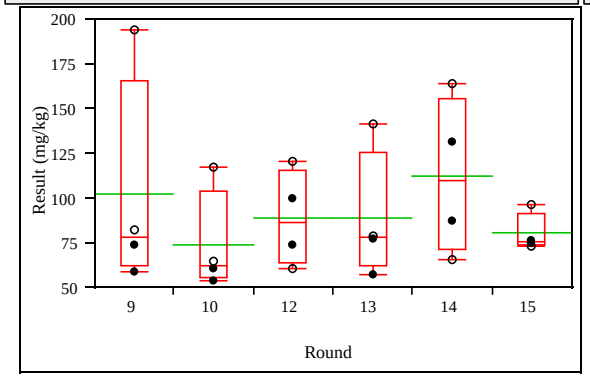
FIGURE 1 (CONTINUED)

BOX PLOTS OF LEAD CONCENTRATIONS BY ROUND AND COMPARISON OF FRONT AND BACK YARDS FOR ALL ROUNDS COMBINED

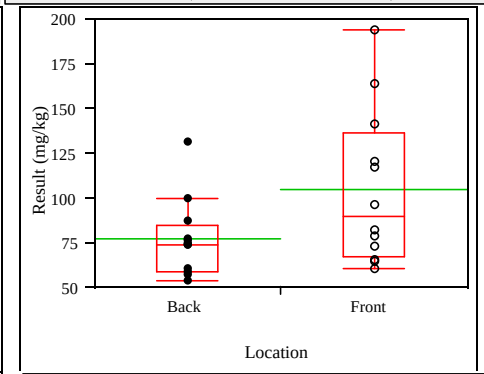
House=24

Open Circles (Front Yards) -- Solid Circles (Back Yards)

All Quadrants Plotted



Back Versus Front (All Rounds Combined)

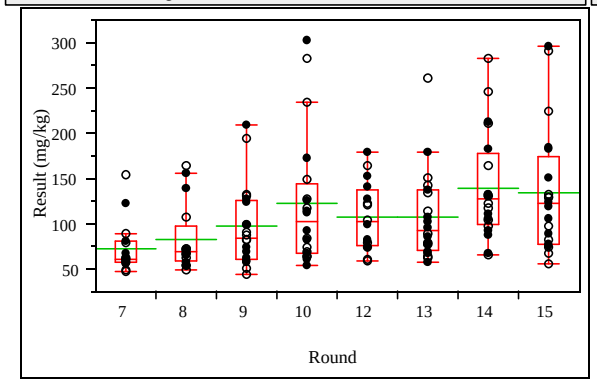


Quantiles

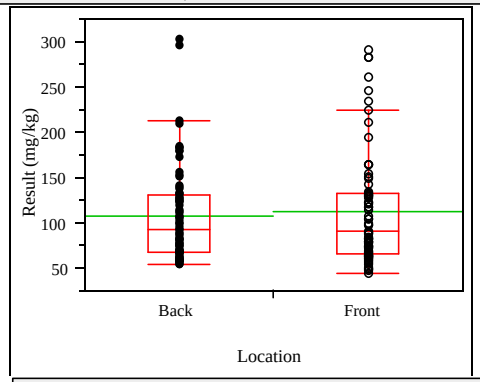
Level	Minimum	10%	25%	Median	75%	90%	Maximum
Back	53	54.2	58.5	73.75	84.5	121.4	131
Front	60	61.2	66.975	89	135.75	184	193

All Houses Within 0.25 Miles of the Smelter Combined

All Houses and Quadrants Plotted



Back Versus Front (All Houses and Rounds Combined)



Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
Back	53	57.5	66.75	91.5	130.25	180.25	302
Front	44	56	65.75	89.5	131	217	290

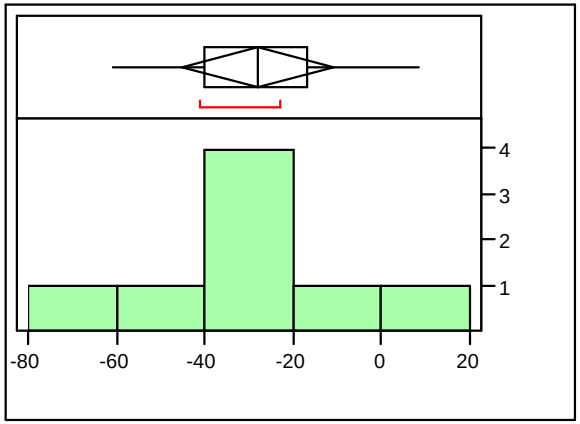
FIGURE 2

RESULTS OF PAIRED-DIFFERENCE TESTS COMPARING LEAD CONCENTRATIONS IN FRONT AND BACK YARDS

House=5

Results of Paired-Difference Tests, H0: Average Difference (Front-Back)= 0

Differences Calculated for the Average Front and Back Yard Concentrations For Each Round



Quantiles

100.0%	maximum	8.50
99.5%		8.50
97.5%		8.50
90.0%		8.50
75.0%	quartile	-16.62
50.0%	median	-27.75
25.0%	quartile	-40.00
10.0%		-61.00
2.5%		-61.00
0.5%		-61.00
0.0%	minimum	-61.00

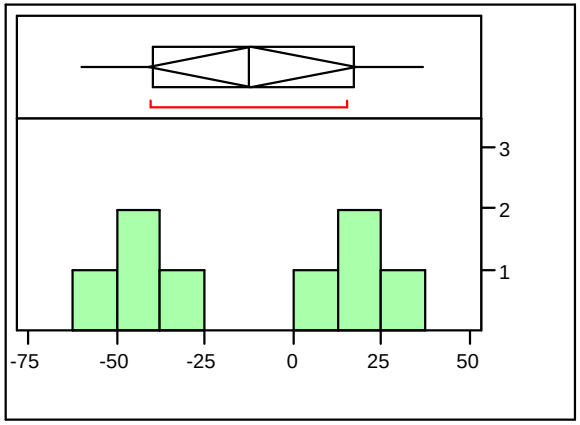
Test Mean=value

Hypothesized Value	0	
Actual Estimate	-27.938	
df	7	
Std Dev	20.332	
	t Test	**Signed-Rank**
Test Statistic	-3.8865	-17.000
Prob > |t|	0.0060	0.016
Prob > t	0.9970	0.992
Prob < t	0.0030	0.008

House=6

Results of Paired-Difference Tests, H0: Average Difference (Front-Back)= 0

Differences Calculated for the Average Front and Back Yard Concentrations For Each Round



Quantiles

100.0%	maximum	36.50
99.5%		36.50
97.5%		36.50
90.0%		36.50
75.0%	quartile	17.25
50.0%	median	-12.25
25.0%	quartile	-39.62
10.0%		-60.00
2.5%		-60.00
0.5%		-60.00
0.0%	minimum	-60.00

Test Mean=value

Hypothesized Value	0	
Actual Estimate	-11.688	
df	7	
Std Dev	35.1161	
	t Test	**Signed-Rank**
Test Statistic	-0.9414	-7.000
Prob > |t|	0.3779	0.383
Prob > t	0.8111	0.809
Prob < t	0.1889	0.191

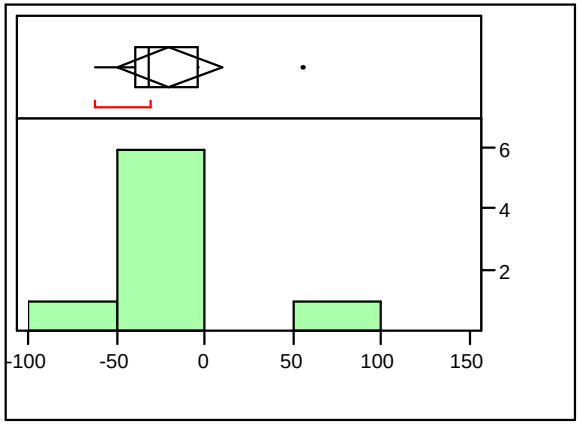
FIGURE 2 (CONTINUED)

RESULTS OF PAIRED-DIFFERENCE TESTS COMPARING LEAD CONCENTRATIONS IN FRONT AND BACK YARDS

House=20

Results of Paired-Difference Tests, H0: Average Difference (Front-Back)= 0

Differences Calculated for the Average Front and Back Yard Concentrations For Each Round



Quantiles

100.0%	maximum	55.50
99.5%		55.50
97.5%		55.50
90.0%		55.50
75.0%	quartile	-4.00
50.0%	median	-32.25
25.0%	quartile	-39.75
10.0%		-62.00
2.5%		-62.00
0.5%		-62.00
0.0%	minimum	-62.00

Test Mean=value

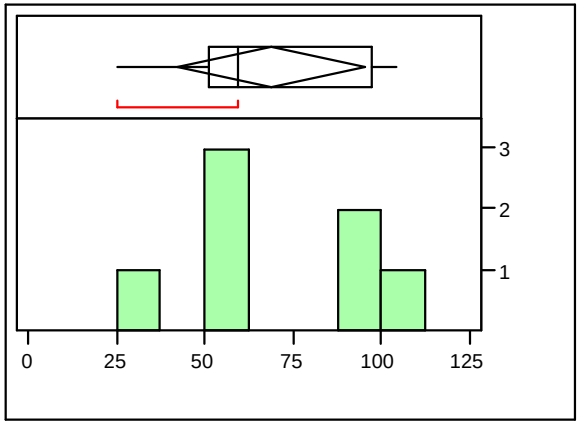
Hypothesized Value	0
Actual Estimate	-19.875
df	7
Std Dev	35.8924

	t Test	Signed-Rank
Test Statistic	-1.5662	-11.000
Prob > |t|	0.1613	0.148
Prob > t	0.9194	0.926
Prob < t	0.0806	0.074

House=22

Results of Paired-Difference Tests, H0: Average Difference (Front-Back)= 0

Differences Calculated for the Average Front and Back Yard Concentrations For Each Round



Quantiles

100.0%	maximum	104.00
99.5%		104.00
97.5%		104.00
90.0%		104.00
75.0%	quartile	97.00
50.0%	median	59.50
25.0%	quartile	51.00
10.0%		25.00
2.5%		25.00
0.5%		25.00
0.0%	minimum	25.00

Test Mean=value

Hypothesized Value	0
Actual Estimate	69
df	6
Std Dev	28.7417

	t Test	Signed-Rank
Test Statistic	6.3516	14.000
Prob > |t|	0.0007	0.016
Prob > t	0.0004	0.008
Prob < t	0.9996	0.992

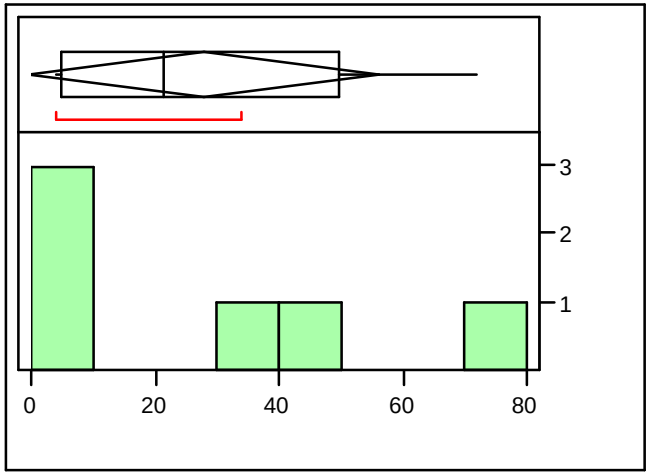
FIGURE 2 (CONTINUED)

RESULTS OF PAIRED-DIFFERENCE TESTS COMPARING LEAD CONCENTRATIONS IN FRONT AND BACK YARDS

House=24

Results of Paired-Difference Tests, H0: Average Difference (Front-Back)= 0

Differences Calculated for the Average Front and Back Yard Concentrations For Each Round



Quantiles

100.0%	maximum	72.000
99.5%		72.000
97.5%		72.000
90.0%		72.000
75.0%	quartile	49.875
50.0%	median	21.550
25.0%	quartile	4.750
10.0%		4.000
2.5%		4.000
0.5%		4.000
0.0%	minimum	4.000

Test Mean=value

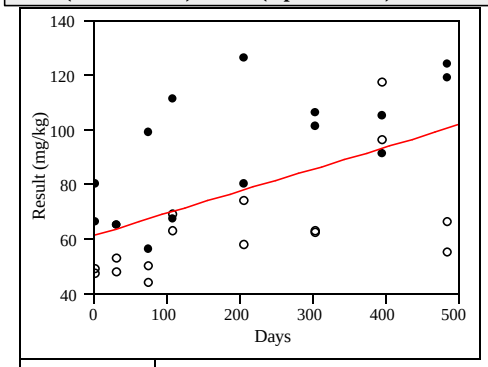
Hypothesized Value	0	
Actual Estimate	27.7667	
df	5	
Std Dev	26.9968	
	t Test	**Signed-Rank**
Test Statistic	2.5193	10.500
Prob > |t|	0.0532	0.031
Prob > t	0.0266	0.016
Prob < t	0.9734	0.984

FIGURE 3

LINEAR REGRESSION ANALYSIS CONDUCTED USING DATA FOR INDIVIDUAL QUADRANTS AND FOR THE MEDIAN OF ALL QUADRANTS

House=5

Back (Solid Circles), Front (Open Circles)



Linear Fit

Result (mg/kg) = 61.09017 + 0.0806629 Days

Summary of Fit

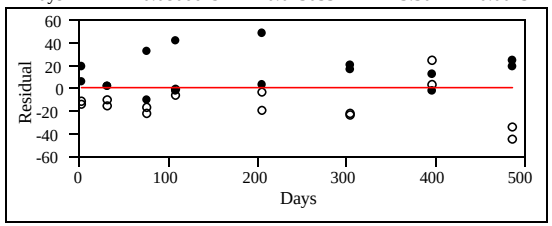
RSquare	0.289832
RSquare Adj	0.26616
Root Mean Square Error	21.71045
Mean of Response	77.34375
Observations (or Sum Wgts)	32

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	5770.908	5770.91	12.2435
Error	30	14140.311	471.34	**Prob > F**
C. Total	31	19911.219		0.0015

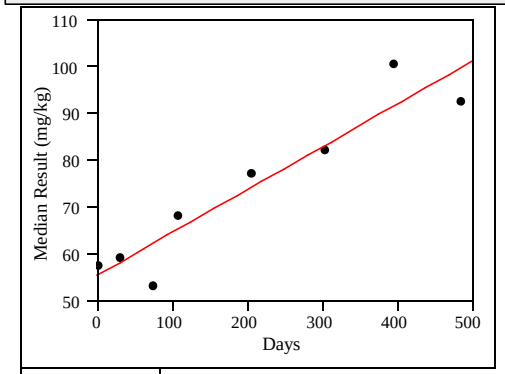
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	61.09017	6.025489	10.14	<.0001
Days	0.0806629	0.023053	3.50	0.0015



House=5

Median of all Quadrants



Linear Fit

Median Result (mg/kg) = 55.277143 + 0.0913665 Days

Summary of Fit

RSquare	0.885244
RSquare Adj	0.866118
Root Mean Square Error	6.323913
Mean of Response	73.6875
Observations (or Sum Wgts)	8

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	1851.0175	1851.02	46.2848
Error	6	239.9513	39.99	**Prob > F**
C. Total	7	2090.9688		0.0005

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	55.277143	3.510261	15.75	<.0001
Days	0.0913665	0.01343	6.80	0.0005

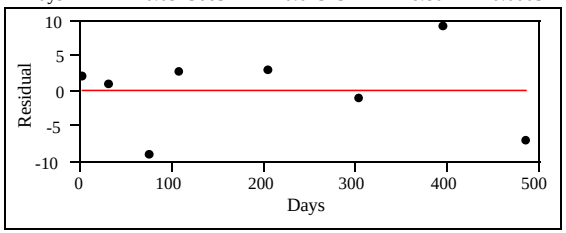
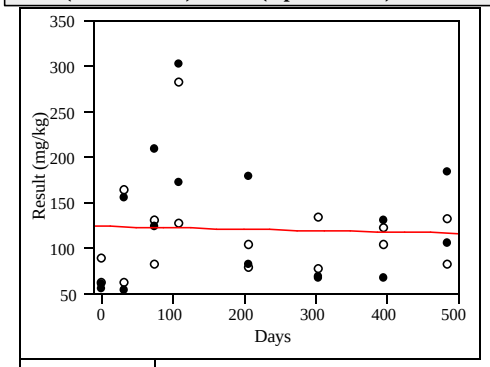


FIGURE 3 (CONTINUED)

LINEAR REGRESSION ANALYSIS CONDUCTED USING DATA FOR INDIVIDUAL QUADRANTS AND FOR THE MEDIAN OF ALL QUADRANTS

House=6

Back (Solid Circles), Front (Open Circles)



Linear Fit

Linear Fit

Result (mg/kg) = 122.75753 - 0.0147982 Days

Summary of Fit

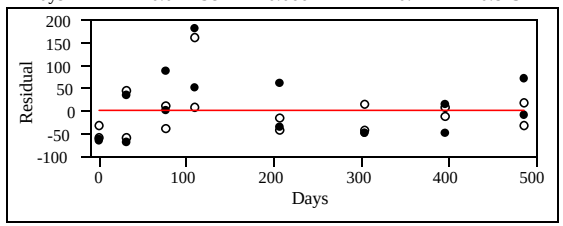
RSquare	0.001652
RSquare Adj	-0.03163
Root Mean Square Error	62.71315
Mean of Response	119.7812
Observations (or Sum Wgts)	32

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	195.28	195.28	0.0497
Error	30	117988.19	3932.94	**Prob > F**
C. Total	31	118183.47		0.8252

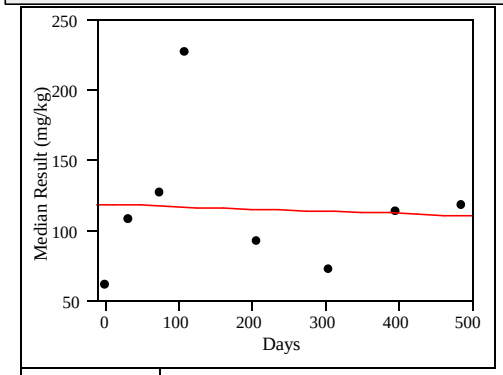
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	122.75753	17.35832	7.07	<.0001
Days	-0.014798	0.066411	-0.22	0.8252



House=6

Median of all Quadrants



Linear Fit

Linear Fit

Median Result (mg/kg) = 118.07292 - 0.0165216 Days

Summary of Fit

RSquare	0.003408
RSquare Adj	-0.16269
Root Mean Square Error	54.45739
Mean of Response	114.75
Observations (or Sum Wgts)	8

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	60.853	60.85	0.0205
Error	6	17793.647	2965.61	**Prob > F**
C. Total	7	17854.500		0.8908

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	118.07292	30.14643	3.92	0.0078
Days	-0.016522	0.115337	-0.14	0.8908

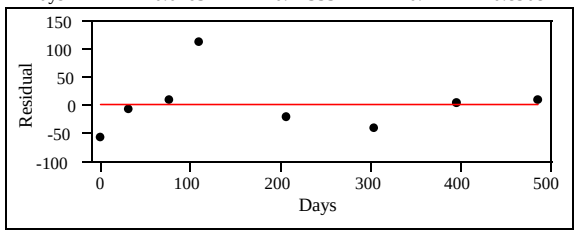
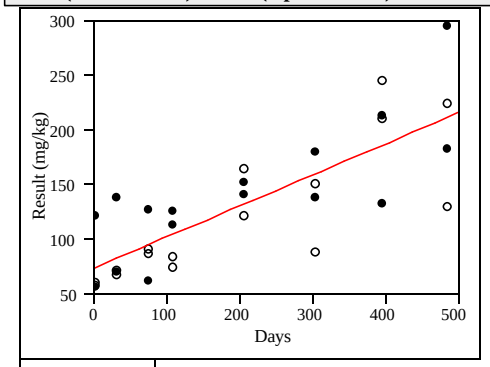


FIGURE 3 (CONTINUED)

LINEAR REGRESSION ANALYSIS CONDUCTED USING DATA FOR INDIVIDUAL QUADRANTS AND FOR THE MEDIAN OF ALL QUADRANTS

House=20

Back (Solid Circles), Front (Open Circles)



Linear Fit

Result (mg/kg) = 72.743316 + 0.283687 Days

Summary of Fit

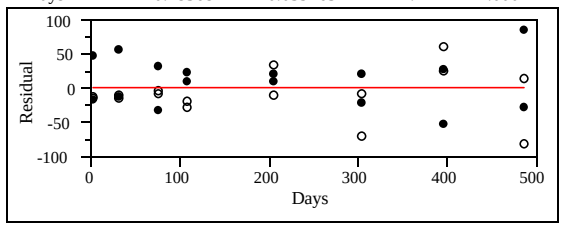
RSquare	0.635033
RSquare Adj	0.622868
Root Mean Square Error	36.97903
Mean of Response	129.9062
Observations (or Sum Wgts)	32

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	71379.77	71379.8	52.1992
Error	30	41023.45	1367.4	**Prob > F**
C. Total	31	112403.22		<.0001

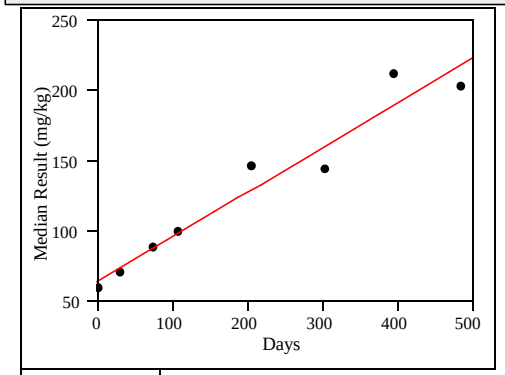
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	72.743316	10.26311	7.09	<.0001
Days	0.283687	0.039265	7.22	<.0001



House=20

Median of all Quadrants



Linear Fit

Median Result (mg/kg) = 63.255978 + 0.3177433 Days

Summary of Fit

RSquare	0.944537
RSquare Adj	0.935293
Root Mean Square Error	14.80167
Mean of Response	127.2812
Observations (or Sum Wgts)	8

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	22386.643	22386.6	102.1804
Error	6	1314.537	219.1	**Prob > F**
C. Total	7	23701.180		<.0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	63.255978	8.216071	7.70	0.0003
Days	0.3177433	0.031433	10.11	<.0001

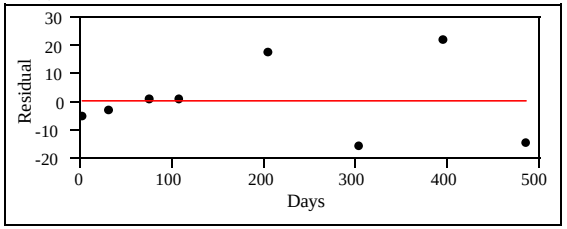
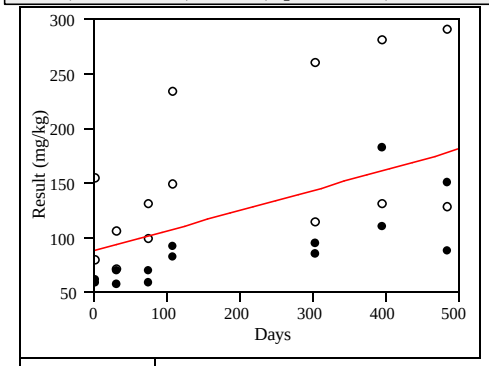


FIGURE 3 (CONTINUED)

LINEAR REGRESSION ANALYSIS CONDUCTED USING DATA FOR INDIVIDUAL QUADRANTS AND FOR THE MEDIAN OF ALL QUADRANTS

House=22

Back (Solid Circles), Front (Open Circles)



Linear Fit

Result (mg/kg) = 87.187288 + 0.1854118 Days

Summary of Fit

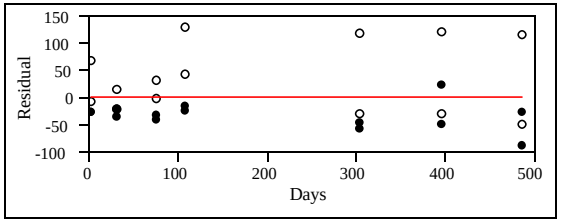
RSquare	0.246292
RSquare Adj	0.217304
Root Mean Square Error	59.90348
Mean of Response	124.4286
Observations (or Sum Wgts)	28

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	30487.76	30487.8	8.4961
Error	26	93299.09	3588.4	**Prob > F**
C. Total	27	123786.86		0.0072

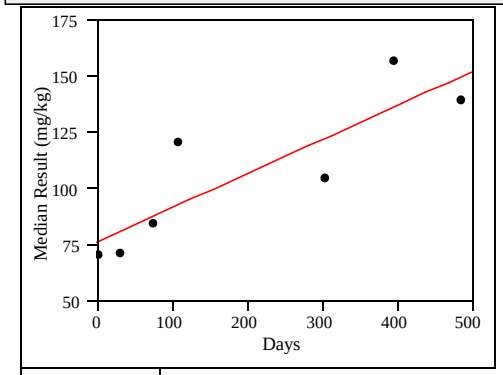
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	87.187288	17.0704	5.11	<.0001
Days	0.1854118	0.06361	2.91	0.0072



House=22

Median of all Quadrants



Linear Fit

Median Result (mg/kg) = 75.855143 + 0.1515036 Days

Summary of Fit

RSquare	0.745437
RSquare Adj	0.694525
Root Mean Square Error	18.6434
Mean of Response	106.2857
Observations (or Sum Wgts)	7

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	5089.0476	5089.05	14.6415
Error	5	1737.8809	347.58	**Prob > F**
C. Total	6	6826.9286		0.0123

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	75.855143	10.62543	7.14	0.0008
Days	0.1515036	0.039594	3.83	0.0123

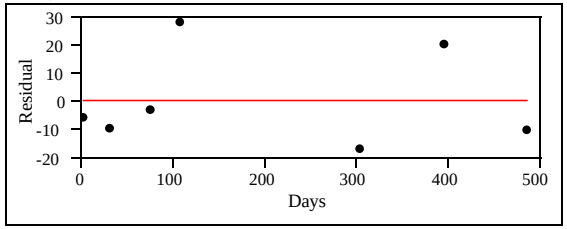
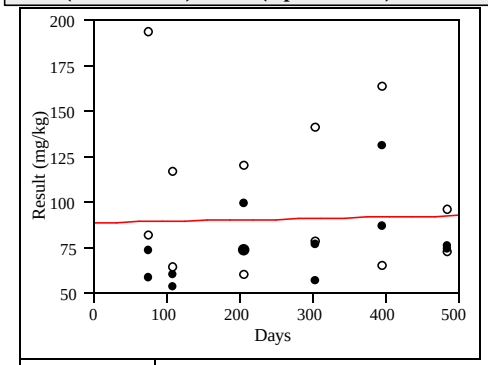


FIGURE 3 (CONTINUED)

LINEAR REGRESSION ANALYSIS CONDUCTED USING DATA FOR INDIVIDUAL QUADRANTS AND FOR THE MEDIAN OF ALL QUADRANTS

House=24

Back (Solid Circles), Front (Open Circles)



Linear Fit

Result (mg/kg) = 88.354788 + 0.0079399 Days

Summary of Fit

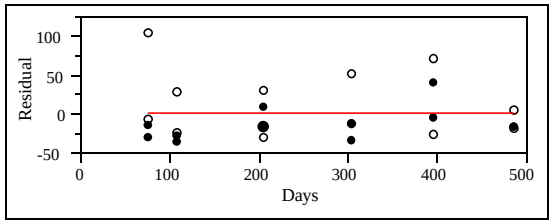
RSquare	0.001093
RSquare Adj	-0.04431
Root Mean Square Error	37.04934
Mean of Response	90.44167
Observations (or Sum Wgts)	24

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	33.043	33.04	0.0241
Error	22	30198.376	1372.65	**Prob > F**
C. Total	23	30231.418		0.8781

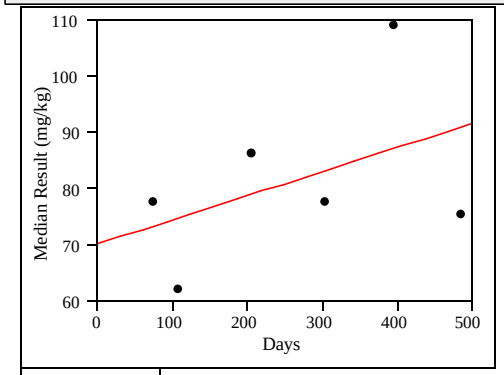
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	88.354788	15.43087	5.73	<.0001
Days	0.0079399	0.051175	0.16	0.8781



House=24

Median of all Quadrants



Linear Fit

Median Result (mg/kg) = 69.959547 + 0.0428616 Days

Summary of Fit

RSquare	0.196329
RSquare Adj	-0.00459
Root Mean Square Error	15.69551
Mean of Response	81.225
Observations (or Sum Wgts)	6

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	240.7225	240.722	0.9772
Error	4	985.3963	246.349	**Prob > F**
C. Total	5	1226.1187		0.3789

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	69.959547	13.07421	5.35	0.0059
Days	0.0428616	0.04336	0.99	0.3789

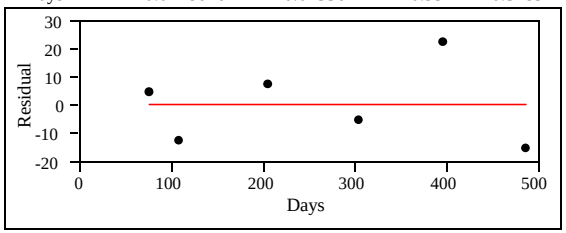
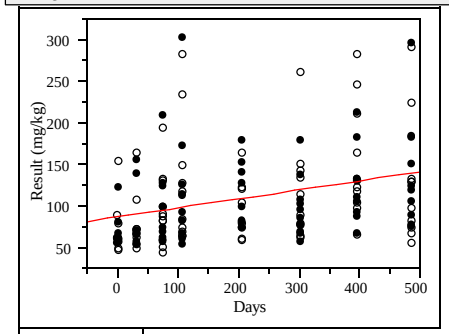


FIGURE 4

REGRESSION ANALYSIS FOR ALL HOUSES POOLED USING DATA FOR INDIVIDUAL QUADRANTS, THE MEDIAN OF ALL QUADRANTS FOR EACH HOUSE, AND THE OVERALL MEDIAN OF ALL QUADRANTS AND HOUSES COMBINED

All Quadrants and Houses Combined



Linear Fit

Result (mg/kg) = 85.977667 + 0.1085883 Days

Summary of Fit

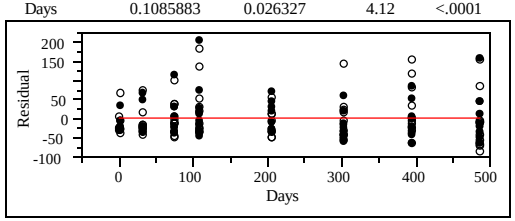
RSquare	0.104361
RSquare Adj	0.098226
Root Mean Square Error	53.65106
Mean of Response	108.9162
Observations (or Sum Wgts)	148

Analysis of Variance

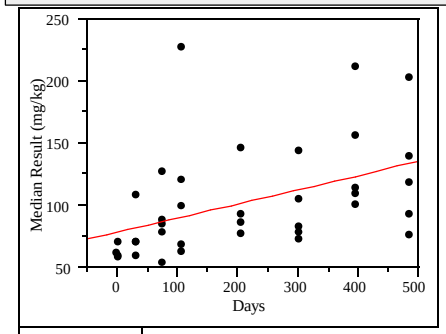
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	48968.12	48968.1	17.0121
Error	146	420251.64	2878.4	**Prob > F**
C. Total	147	469219.76		<.0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	85.977667	7.097784	12.11	<.0001
Days	0.1085883	0.026327	4.12	<.0001



Medians Plotted for Individual Houses for Each Date



Linear Fit

Median Result (mg/kg) = 77.391699 + 0.1143305 Days

Summary of Fit

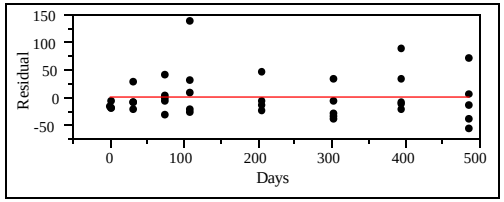
RSquare	0.201839
RSquare Adj	0.179035
Root Mean Square Error	39.1574
Mean of Response	101.5432
Observations (or Sum Wgts)	37

Analysis of Variance

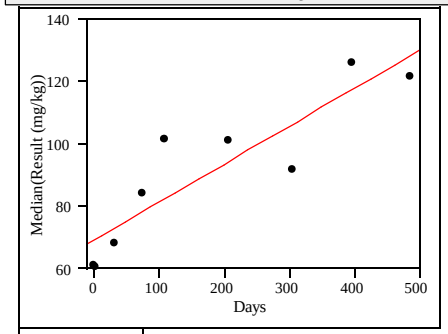
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	13570.985	13571.0	8.8508
Error	35	53665.581	1533.3	**Prob > F**
C. Total	36	67236.566		0.0053

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	77.391699	10.36068	7.47	<.0001
Days	0.1143305	0.03843	2.98	0.0053



Grand Median of all Houses and Quadrants Plotted



Linear Fit

Median(Result (mg/kg)) = 68.744595 + 0.1217734 Days

Summary of Fit

RSquare	0.801147
RSquare Adj	0.772739
Root Mean Square Error	11.64334
Mean of Response	90.55556
Observations (or Sum Wgts)	9

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	3823.2503	3823.25	28.2018
Error	7	948.9719	135.57	**Prob > F**
C. Total	8	4772.2222		0.0011

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	68.744595	5.650786	12.17	<.0001
Days	0.1217734	0.022931	5.31	0.0011